doi:10.3772/j.issn.1006-6748.2014.03.008

Modified Levenberg-Marquardt algorithm for source localization using AOAs in the presence of sensor location errors①

Wu Xinhui (吴鑫辉)②[*,**], Huang Gaoming[*], Gao Jun[*]
([*] College of Electronic Engineering, Naval University of Engineering, Wuhan 430034, P. R. China)
([**] Automatization of Command Institute, Academy of Navy Equipment, Beijing 100036)

Abstract

In this paper, by utilizing the angle of arrivals (AOAs) and imprecise positions of the sensors, a novel modified Levenberg-Marquardt algorithm to solve the source localization problem is proposed. Conventional source localization algorithms, like Gauss-Newton algorithm and Conjugate gradient algorithm are subjected to the problems of local minima and good initial guess. This paper presents a new optimization technique to find the descent directions to avoid divergence, and a trust region method is introduced to accelerate the convergence rate. Compared with conventional methods, the new algorithm offers increased stability and is more robust, allowing for stronger non-linearity and wider convergence field to be identified. Simulation results demonstrate that the proposed algorithm improves the typical methods in both speed and robustness, and is able to avoid local minima.

Key words: source localization, angle of arrivals (AOAs), nonlinear least-squares estimators, Levenberg-Marquardt algorithm

0 Introduction

Source localization by using the angle-of-arrivals (AOAs) measurements has a wide application in areas such as localization of mobile phones, radar, sonar and electronic warfare. Various algorithms have been proposed to estimate the source location[1-3]. Previous methods assume that the sensor positions are known exactly, which may not be true in practice. Recently, Ho, et al[2] proposed a closed-form solution to estimate the source location using time difference of arrival (TDOA) measurements when the sensor location errors are presented. Lu, et al[3] employed the Taylor-series technique using AOAs measurements to find the source location when sensor positions have errors, which indicates that these methods improve the source location accuracy significantly when sensor position errors are taken into account[6].

It requires solving a set of nonlinear observe equations to find the position of an emitter by using AOAs measurements. The existing solutions to this problem can be loosely divided into two categories: Some methods are nonlinear least-squares (NLS) estimators[3,4], the others are closed-form solutions[5]. Both of these methods are able to reach the CRLB accuracy. The closed-form solutions are obtained from linearization of the set of nonlinear equations by transforming the unknown emitter location vector[6]. However, the "linearization" process may depend on the special position parameters, which is not an all-purpose method. The NLS estimator is a usual method to solve the sets of nonlinear equations but requires a good initial guess. The Gaussian-Newton method[3] is the most common NLS estimator employed to estimate the emitter position. Unfortunately, there exists three problems in the Gaussian-Newton method for emitter localization: (1) the Gauss-Newton method fails for strongly non-linear equation problems; (2) it requires a good initial position guess, which is used to avoid local minima and divergence; (3) the convergence of the iterative process is not guaranteed in practice. In Ref. [7], a technique based on grid-search was developed to find the suitable initial guess, however, the grid-search computation is expensive when the emitter source is unknown.

The Levenberg-Margquardt (LM) algorithm is an effective nonlinear least-squares method that combines the robustness of steepest descent method and the computational efficiency of the Gauss-Newton method. Nonlinear least squares problems have been successfully solved by the various implementations of the LM method[8-10]. The LM method is much less dependent

① Supported by the National High Technology Research and Development Programme of China (No. 2011AA7014061).
② To whom correspondence should be addressed. E-mail: wuxinhui009@163.com
Received on Oct. 23, 2012

on the accuracy of the initial guess and more robust than previous methods, thus the stability and efficiency of nonlinear localization estimation can be notably improved. However, compared with other iterated least square methods, the standard LM method may suffer local minima and slower convergence problems.

This paper presents an accelerated Levenberg-Marquardt method for the application of AOAs localization. It works with combination of the Gauss-Newton and the LM directions so as to force the several nonlinear localization equations to converge no matter whether the initial guess is at the solution's neighborhood or not. A radius of trust region is used to control the size of step which is capable of accelerating the convergence rate. Moreover, a stopping criterion is constructed to indicate whether the iteration converges to the global minimal value. With the use of the LM algorithm and the trust regions, the method converges from a wider range of parameters without declining the estimation precision.

This paper is organized as follows: Section 1 presents the measurement model for passive location in the presence of sensor location errors. In Section 2, the modified Levenberg-Margquardt method for source localization is developed. Simulations are included in Section 3 to evaluate the estimator performance by comparing it with Gauss-Newton and standard LM methods. Finally, the conclusions are drawn in Section 4.

1 AOAs location model

Fig. 1 gives an illustration of a typical operational scenario where the AOAs-based geolocation technique is used. Here 2-D source localization is considered with the sensors have random position errors. Let $\boldsymbol{s}_i = [x_i, y_i]^{\mathrm{T}}$, $i = 1, 2, \cdots, M$ be the available sensor locations with position error. The true sensor positions are $\boldsymbol{s}_i^o = [x_i^o, y_i^o]^{\mathrm{T}}$, and the true source position is denoted as $\boldsymbol{u}^o = [x^o, y^o]^{\mathrm{T}}$, $i = 1, 2, \cdots, M$. The AOAs measurements between the source and sensor i is

$$\boldsymbol{\beta}^{\mathrm{T}} = [\beta_1, \cdots, \beta_M]^{\mathrm{T}} = \left[\tan^{-1}\left(\frac{y^o - y_1^o}{x^o - x_1^o}\right), \cdots, \tan^{-1}\left(\frac{y^o - y_M^o}{x^o - x_M^o}\right)\right]^{\mathrm{T}} + \boldsymbol{n} \tag{1}$$

The AOA measurement noise is denoted as $\boldsymbol{n} = [\Delta\beta_1, \Delta\beta_2, \cdots, \Delta\beta_M]^{\mathrm{T}}$, $\boldsymbol{n}$ is assumed to be a zero mean Gaussian vector with covariance matrix $\mathrm{E}[\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}] = \boldsymbol{Q}_n$.

The available sensor positions with unknown noise can be represented as:

$$\boldsymbol{s}^{\mathrm{T}} = [\boldsymbol{s}_1, \cdots, \boldsymbol{s}_M]^{\mathrm{T}} = [\boldsymbol{s}_1^o, \cdots, \boldsymbol{s}_M^o]^{\mathrm{T}} + \boldsymbol{m} \tag{2}$$

where the covariance matrix for vector $\boldsymbol{m}$ is $\boldsymbol{Q}_m = \mathrm{E}(\boldsymbol{m}\boldsymbol{m}^{\mathrm{T}})$. $\boldsymbol{m} = [\Delta s_1; \Delta s_2; \cdots, \Delta s_M]$ is the vector of sensor position errors, and it is assumed to be independent of the measurement noise $\boldsymbol{n}$.

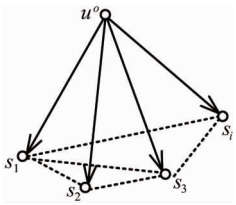

Fig. 1 Two-dimensional AOA emitter localization geometry with i receviers

Let $\boldsymbol{p} = [\boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{s}^{\mathrm{T}}]^{\mathrm{T}}$ represent the vector of the AOA measurements and available sensor positions, $\boldsymbol{w} = [(\boldsymbol{u}^o)^{\mathrm{T}}, (\boldsymbol{s}_1^o)^{\mathrm{T}}, \cdots, (\boldsymbol{s}_M^o)^{\mathrm{T}}]^{\mathrm{T}}$ be the unknown parameter vector of the true source and sensor locations, and $\boldsymbol{e} = [\boldsymbol{n}^{\mathrm{T}}, \boldsymbol{m}^{\mathrm{T}}]^{\mathrm{T}}$ be the vector of measurement error and sensor position error. With the AoAs measurements in Eq. (1) and sensor positions in Eq. (2), the mathematical problem of estimating $\boldsymbol{w}$ can be expressed as the following algebraic relation:

$$\boldsymbol{p} = f(\boldsymbol{w}) + \boldsymbol{e} \tag{3}$$

the measurement error covariance matrix $\sum$ is given as:

$$\sum = E(\boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}) = \begin{bmatrix} \boldsymbol{Q}_n & \\ & \boldsymbol{Q}_m \end{bmatrix} \tag{4}$$

The nonlinear least squares cost function for estimating the unknown parameter vector $\boldsymbol{w}$ in the azimuth plane is given as:

$$\hat{\boldsymbol{w}}_p = \arg\min_{w} \boldsymbol{J}_{\mathrm{NLS}} = \arg\min_{w} \| \boldsymbol{p} - \mathrm{f}(\boldsymbol{w}) \|_W^2 \tag{5}$$

where $\| \cdot \|_W = \sqrt{(\cdot)^T \sum^{-1} (\cdot)}$ denotes a weighted norm, with the error covariance matrix $\sum$.

2 The accelerated Levenberg-Marquardt location algorithm

2.1 Conventional methods

The conventional method needs the initial guess $\boldsymbol{w}_g = [\boldsymbol{u}_g^{\mathrm{T}}, \boldsymbol{s}_g^{\mathrm{T}}]^{\mathrm{T}}$ closing to the true parameter vector $\boldsymbol{w}$. For small $\boldsymbol{\delta w}_g$, expand $f(\boldsymbol{w})$ around $\boldsymbol{w}_g$ through Taylor series and keep only up to the first order term

$$f(\boldsymbol{w}) \approx \mathrm{f}(\boldsymbol{w})|_{w_g} + \boldsymbol{F}\boldsymbol{\delta w}_g \tag{6}$$

where $\boldsymbol{F}$ is a matrix equal to the partial derivative of $\mathrm{f}(\boldsymbol{w})$ with respect to $\boldsymbol{w}$. Inserting Eq. (6) in Eq. (5), the output of cost function is

$$\boldsymbol{J}_{\mathrm{NLS}} \approx (\boldsymbol{p} - f(\boldsymbol{w})|_{w_g} - \boldsymbol{F}\boldsymbol{\delta w}_g)^{\mathrm{T}} \sum^{-1} (\boldsymbol{p} - f(\boldsymbol{w})|_{w_g} - \boldsymbol{F}\boldsymbol{\delta w}_g) \tag{7}$$

Let $\boldsymbol{h} = \boldsymbol{p} - f(\boldsymbol{w})\mid_{w_g}$, we can express Eq. (7) as

$$\begin{aligned}\boldsymbol{J}_{\mathrm{NLS}} &\approx (\boldsymbol{h} - \boldsymbol{F}\delta\boldsymbol{w}_g)^{\mathrm{T}} \sum^{-1} (\boldsymbol{h} - \boldsymbol{F}\delta\boldsymbol{w}_g) \\ &= \boldsymbol{h}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} - 2(\delta\boldsymbol{w}_g)^{\mathrm{T}} \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} \\ &\quad + (\delta\boldsymbol{w}_g)^{\mathrm{T}} \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F}(\delta\boldsymbol{w}_g) \end{aligned} \tag{8}$$

The Gauss-Newton step $\delta\boldsymbol{w}_g$ minimizes $\boldsymbol{J}_{\mathrm{NLS}}$, and it is obvious that the gradient $\boldsymbol{J}'_{\mathrm{NLS}}$ and the Hessian of $\boldsymbol{J}_{NLS}$ are derived as:

$$\begin{aligned}\frac{\partial \boldsymbol{J}_{\mathrm{NLS}}}{\partial(\delta\boldsymbol{w}_g)} &= \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F}(\delta\boldsymbol{w}_g) - 2\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h}, \\ \frac{\partial^2 \boldsymbol{J}_{\mathrm{NLS}}}{\partial(\delta\boldsymbol{w}_g)^2} &= \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F} \end{aligned} \tag{9}$$

where $\boldsymbol{F}$ is derived from Eq. (6), $\boldsymbol{J}''_{\mathrm{NLS}}$ is independent of $\delta\boldsymbol{w}_g$. If $\boldsymbol{F}$ has full rank, $\boldsymbol{J}''_{\mathrm{NLS}}$ is positive definite. From Ref. [11], this implies that $\boldsymbol{J}_{\mathrm{NLS}}$ has a unique minimizer, which can be found by solving

$$\delta\boldsymbol{w}_g = (\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F})^{-1} \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} \tag{10}$$

To obtain the solution, it is needed to compute $\delta\boldsymbol{w}_g$ by Eq. (10) at each iteration, replace

$$\hat{\boldsymbol{w}}_{k+1} = \hat{\boldsymbol{w}}_k + \alpha(\delta\boldsymbol{w}_g) \tag{11}$$

The classical Gauss-Newton method uses $\alpha = 1$ in all steps. The method with line search can be shown to have guaranteed convergence, provided that

(1) $\{\boldsymbol{w} \mid \mathrm{f}(\boldsymbol{w}) \leqslant \mathrm{f}(\boldsymbol{w}_0)\}$ is bounded,

(2) the Jacobian matrix $\boldsymbol{F}$ has full rank in all steps.

The standard Gauss-Newton scheme has several hidden drawbacks. Firstly, Because the function $\mathrm{f}(\boldsymbol{w})$ is nonlinear and the relativity among the portions of $f(\boldsymbol{w})$ always exists, it may happen that the gradient matrix in Eq. (10) at some test points degenerates. When $(\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F})^{-1}$ is singular, it jumps out of the iteration without reaching the minimum. Secondly, the Gauss-Newton method may diverge or converge to a saddle point or a point of local minimum. This would require good initial guesses for the parameters which are in practice not available.

2.2 Accelerated LM (A-LM) algorithm

The LM algorithm is an iterative technique that locates the minimum of a multivariate function expressed as the sum of squares of non-linear real-valued functions. It gives a good compromise between the speed of the Newton algorithm and the stability of the steepest descent method[8]. With the use of a damped parameter μ, the step $\delta\boldsymbol{w}_l$ is defined by the following modification to Eq. (10),

$$(\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F} + \mu\boldsymbol{I})\delta\boldsymbol{w}_l = 2\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} \tag{12}$$

And the step $\delta\boldsymbol{w}_l$ in each step can be calculated as

$$\delta\boldsymbol{w}_l = 2(\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F} + \mu\boldsymbol{I})^{-1} \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} \tag{13}$$

Thus the solution for the each step is

$$\begin{cases} \hat{\boldsymbol{w}}_{k+1} = \hat{\boldsymbol{w}}_k + \delta\boldsymbol{w}_l \\ \delta\boldsymbol{w}_l = 2(\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F} + \mu\boldsymbol{I})^{-1} \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} \end{cases} \tag{14}$$

From Eq. (14), the damping parameter μ influences both the direction and the size of the step.

2.2.1 comments on the parameter μ

It is known that the Gauss-Newton method is good at the final stage of the minimization process, but it is better to use the gradient method for the first iterations. The damping parameter μ has several effects:

(1) For all $\mu > 0$ the coefficient matrix is positive definite, and this ensures that the inverse matrix $(\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F} + \mu\boldsymbol{I})^{-1}$ exists.

(2) For large values of μ we get

$$\delta\boldsymbol{w}_l \approx \frac{1}{\mu}\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} \tag{15}$$

which is a short step in the steepest descent direction. It is good if the current iteration is far from the solution.

(3) If μ is very small, then $\delta\boldsymbol{w}_l \approx \delta\boldsymbol{w}_g$, which is a good step in the final stages of the iteration, where the $\hat{\boldsymbol{w}}_k$ is close to the real position $\boldsymbol{w}$.

2.2.2 the control of optimal iteration step

It is necessary to control the step size carefully once the direction of the correction vector has been established. In view of LM method, the LM method leads to slower convergence than the Gauss-Newton. The trust region could be used to find an optimal approximate step which is a combination of Gauss-Newton step and LM step.

There are two candidates for the step to take from the current point $\hat{\boldsymbol{w}}_k$: the LM step $\delta\boldsymbol{w}_l$ and the Gauss-Newton step $\delta\boldsymbol{w}_g$. When the trust region has radius Δ, the strategy for choosing the step $\delta\boldsymbol{w}_{gl}$ can be illustrated in Fig. 2.

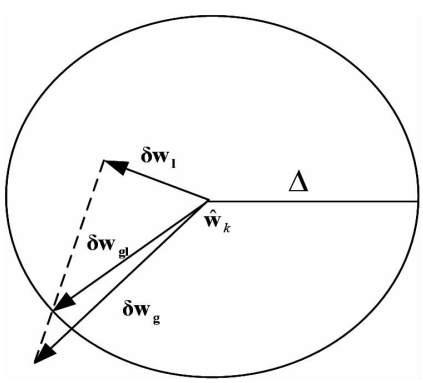

Fig. 2 The radius of trust region and the size of step

$$\boldsymbol{\delta w}_{gl} = \begin{cases} \boldsymbol{\delta w}_g, & \| \boldsymbol{\delta w}_g \| \leqslant \Delta \\ \boldsymbol{\delta w}_l + \beta \times (\boldsymbol{\delta w}_g - \boldsymbol{\delta w}_l), & \| \boldsymbol{\delta w}_g \| > \Delta \end{cases} \tag{16}$$

The parameter β in Eq. (16) is unknown, which is used to control the iterative steps. We consider the iterative step is smaller than the trust region radius Δ, then the optimal size of the step $\| \boldsymbol{\delta w}_{gl} \| = \Delta$. Thus it is obvious that the parameter β satisfy the following relation

$$\begin{aligned} f(\beta) &= \| \boldsymbol{\delta w}_l + \beta(\boldsymbol{\delta w}_g - \boldsymbol{\delta w}_l) \|^2 - \Delta^2 \\ &= \| (\boldsymbol{\delta w}_g - \boldsymbol{\delta w}_l) \|^2 \beta^2 + 2(\boldsymbol{\delta w}_l)^{\mathrm{T}} \\ &\quad (\boldsymbol{\delta w}_g - \boldsymbol{\delta w}_l)\beta + \| \boldsymbol{\delta w}_l \| - \Delta^2 = 0 \end{aligned} \tag{17}$$

The root for the second degree polynomial Eq. (17) exists, because $f(0) = \| \boldsymbol{\delta w}_l \| - \Delta^2 < 0$ and $f(1) = \| \boldsymbol{\delta w}_g \| - \Delta^2 > 0$. Thus, Eq. (17) has one negative root and one root in $[0,1]$. The most accurate computation of β is given by

$$\beta = \begin{cases} (-c + sqrt(c^2 + \| (\boldsymbol{\delta w}_g - \boldsymbol{\delta w}_l) \|^2 \\ \quad (\Delta^2 - \| \boldsymbol{\delta w}_l \|^2))) / \| (\boldsymbol{\delta w}_g - \boldsymbol{\delta w}_l) \|^2 & c \leqslant 0 \\ (\Delta^2 - \| \boldsymbol{\delta w}_l \|^2) / (c + sqrt(c^2 + \| (\boldsymbol{\delta w}_g - \boldsymbol{\delta w}_l) \|^2 \\ \quad (\Delta^2 - \| \boldsymbol{\delta w}_l \|^2))) & c > 0 \end{cases} \tag{18}$$

where $c = (\boldsymbol{\delta w}_l)^{\mathrm{T}} (\boldsymbol{\delta w}_g - \boldsymbol{\delta w}_l)$.

During iteration the radius Δ of trust region is controlled by the gain ratio ρ, which can represent the quality of model

$$\rho = \frac{F(\hat{\boldsymbol{w}}_k) - F(\hat{\boldsymbol{w}}_k + \boldsymbol{\delta w}_{gl})}{L(0) - L(\boldsymbol{\delta w}_{gl})} \tag{19}$$

where $F(\boldsymbol{w}) = \frac{1}{2} \| f(\boldsymbol{w}) \|^2$, the denominator $L(\boldsymbol{\delta w}) = \frac{1}{2} \| f(\boldsymbol{w}) |_{w_g} + \boldsymbol{F}\boldsymbol{\delta w} \|^2$ is the linear model from Eq. (6). Thus it is easy to get that a large value of ρ indicates that the linear model is good, and we can increase Δ so that the next iteration step is closer to the Gauss-Newton step. If ρ is small, then the $L(\boldsymbol{\delta w})$ is a poor approximation, and we should decrease Δ with the twofold aim of getting closer to the LM direction and reducing the step length.

Corresponding to the two cases in Eq. (16), it is easy to compute the denominator in Eq. (19):

$$\begin{aligned} &L(0) - L(\boldsymbol{\delta w}_{gl}) \\ &= 0.5 \| f(\hat{\boldsymbol{w}}_k) \|^2 - 0.5 \| f(\hat{\boldsymbol{w}}_k) + \boldsymbol{F}\boldsymbol{\delta w}_{gl} \|^2 \\ &= 0.5 [f^{\mathrm{T}}(\hat{\boldsymbol{w}}_k) f(\hat{\boldsymbol{w}}_k) - (f(\hat{\boldsymbol{w}}_k) + \boldsymbol{F}\boldsymbol{\delta w}_{gl})^{\mathrm{T}} (f(\hat{\boldsymbol{w}}_k) \\ &\quad + \boldsymbol{F}\boldsymbol{\delta w}_{gl})] \\ &= -\boldsymbol{\delta w}_{gl}^{\mathrm{T}} \boldsymbol{F}^{\mathrm{T}} f(\hat{\boldsymbol{w}}_k) - 0.5 (\boldsymbol{\delta w}_{gl}^{\mathrm{T}} \boldsymbol{F}^{\mathrm{T}} \boldsymbol{F} \boldsymbol{\delta w}_{gl}) \end{aligned} \tag{20}$$

Note that the $\boldsymbol{\delta w}_{gl}$ is the iteration step at time $k+1$, while $\hat{\boldsymbol{w}}_k$ is the new estimation vector at time k. The gain ratio ρ is computed at each step to control the size of radius Δ.

2.2.3 the stop criterion for the global minima

For strongly non-linear problems, the cost function Eq. (5) may have multiple minima. In such cases, the Newton type algorithms would have the well-known problem of local convergence. An effective stopping criterion should help the reflection of a global minimum. The conventional stopping criterion proposed in Ref. [3] $\| \boldsymbol{\delta w} \| < \varepsilon$ or $\boldsymbol{J}'_{NLS} < \varepsilon$ may converge to local minima.

To avoid local minima, a combination of LM method and a random search based on properties of the AOAs localization is introduced. The global minimum must satisfy the following two conditions: (1) the value of μ is almost 0 if w matches the minimum. In the section 2.2.1 we have proved that when the iteration value is near the minimum, the damped parameter $\mu \to 0$; (2) the function $E(\boldsymbol{w}) = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}$ is sufficiently small near the true source localization. From Eq. (3) we can see that $\boldsymbol{e} = \boldsymbol{p} - \mathrm{f}(\boldsymbol{w})$, where $\boldsymbol{p}$ is the measurement vector and $\mathrm{f}(\boldsymbol{w})$ is the estimation vector. When the estimates are near the true source location, the $\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}$ should be very small because the formation of observation error is small, and sensor position error is also small. This can be proved by simulation (Fig. 4) in section 4.

In order not to disturb the line search method from converging to the desired solution by random jumps, those jumps are only allowed if the absolute values could not satisfy the condition (1) and condition (2) simultaneously. Now we can formulate our stopping criterion and random jump when the value of μ has to be smaller than the bound $\mu_{\min}$ but the $E(\boldsymbol{w})$ is larger than the lower bound $E_{\min}$:

$$\begin{aligned} &\text{for step } k = 1 \quad \text{to} \quad n \text{ do} \\ &\quad \hat{\boldsymbol{w}}_{k+1} = \hat{\boldsymbol{w}}_k + \boldsymbol{\delta w}_{gl}; \\ &\quad \text{if } \mu < \mu_{\min} \text{ and } E(\hat{\boldsymbol{w}}_{k+1}) > E_{\min} \\ &\qquad \text{Then } \hat{\boldsymbol{w}}_{k+1} = \hat{\boldsymbol{w}}_{k+1} = \hat{\boldsymbol{w}}_{k+1} + \mathrm{rand}(\boldsymbol{w}); \\ &\quad \text{end if} \\ &\quad \text{end for} \end{aligned} \tag{21}$$

$\mathrm{rand}(\boldsymbol{w})$ is the random variables uniformly distributed over the parameter range of $\boldsymbol{w}$.

In Table 1, the updating damping factor μ and the trust region radius Δ are both controlled by the gain ratio ρ. A small value of ρ indicates that the iterative step is not good and the damping factor μ should increase and the trust region radius Δ should decrease. If the ρ is large, which means the iterative step is a good approximation to the model, and we should decrease μ and increase Δ. The thresholds of gain ratio are set to $\rho = 0.25$ or $\rho = 0.75$ because the updating of parameters

change across the thresholds 0.25 and 0.75 is slightly fluctuant, and we can slow down convergence[11].

Table 1 Algortihme for accelated levenberg-marquardt method

Input: the AOA measurements and available sensor positions vector $\boldsymbol{p} = [\boldsymbol{\beta}^{\mathrm{T}}, \mathbf{s}^{\mathrm{T}}]^{\mathrm{T}}$. A random initial parameter estimate $\mathbf{w}_0 \in R$, the stopping value $\varepsilon_1 = \varepsilon_2 = 10^{-8}$, the damping parameter μ: $= \tau \times \max\{\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}\}$, the radius of trust region $\Delta = \Delta_0$.

Output: a vector $\boldsymbol{w}_k \in R$ minimizing Eq. (4).

Begin: stop: = ($\| \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} \|_{\infty} \leqslant \varepsilon_1$)

while (not stop) and ($k < k_{\max}$)

k: = $k + 1$; $\boldsymbol{h} = \boldsymbol{p} - f(\boldsymbol{w})|_{w_k}$;

Slove $(\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F} + \mu \boldsymbol{I})\delta \boldsymbol{w}_l = 2\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h}$;

Slove $(\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{F})\delta \boldsymbol{w}_g = 2\boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h}$;

Compute $\delta \boldsymbol{w}_{gl}$ from Eq. (15);

if $\| \delta \boldsymbol{w}_l \| \leqslant \varepsilon_2 (\| \hat{\boldsymbol{w}}_k \| + \varepsilon_2)$

Stop: = true;

else

$\hat{\boldsymbol{w}}_{\mathrm{new}} = \hat{\boldsymbol{w}}_k + \delta \boldsymbol{w}_{gl}$;

ρ: $= \dfrac{\| F(\hat{\boldsymbol{w}}_{\mathrm{new}}) \| - \| F(\hat{\boldsymbol{w}}_{\mathrm{new}} + \delta \boldsymbol{w}_{gl}) \|}{\mathrm{L}(0) - \mathrm{L}(\delta \boldsymbol{w}_{gl})}$;

if $\rho > 0.75$

$\hat{\boldsymbol{w}}_{k+1} = \boldsymbol{w}_{\mathrm{new}}$;

stop: = ($\| \boldsymbol{F}^{\mathrm{T}} \sum^{-1} \boldsymbol{h} \|_{\infty} \leqslant \varepsilon_1$);

μ: $= \mu \times \max\{1/3, 1 - (2\rho - 1)^3\}$; v: = 2;

$\Delta = \max\{\Delta, 3 \times \| \delta \boldsymbol{w}_{gl} \|\}$;

else if $0 < \rho < 0.25$

$\hat{\boldsymbol{w}}_{k+1} = \boldsymbol{w}_{\mathrm{new}}$;

stop: = ($\Delta \leqslant \varepsilon_2 (\| \hat{\boldsymbol{w}}_{k+1} \| + \varepsilon_2)$);

μ: $= \mu \times \max\{1/3, 1 - (2\rho - 1)^3\}$; v: = 2;

$\Delta = \Delta/2$;

else

μ: $= \mu \times v$; v: $= 2 \times v$;

$\Delta = \Delta/2$;

end if

$E(\boldsymbol{w}_{k+1}) = (\boldsymbol{p} - f(\boldsymbol{w}_{k+1}))^{\mathrm{T}}(\boldsymbol{p} - f(\boldsymbol{w}_{k+1}))$;

if $\mu < \mu_{\min}$ and $E(\hat{\boldsymbol{w}}_{k+1}) > E_{\min}$

Then $\hat{\boldsymbol{w}}_{k+1} = \hat{\boldsymbol{w}}_{k+1} + \mathrm{rand}(\boldsymbol{w})$;

end if

end if

end while

3 Simulation results

This section provides a performance comparison of the proposed method with the Gauss-Newton method and the CRLB in Ref. [3]. The sensor positions used for simulation are $[-650\mathrm{m}, -493\mathrm{m}]^{\mathrm{T}}$, $[444\mathrm{m}, 562\mathrm{m}]^{\mathrm{T}}$, $[860\mathrm{m}, -843\mathrm{m}]^{\mathrm{T}}$, $[-515\mathrm{m}, 191\mathrm{m}]^{\mathrm{T}}$, $[222\mathrm{m}, -150\mathrm{m}]^{\mathrm{T}}$, $[180\mathrm{m}, 340\mathrm{m}]^{\mathrm{T}}$. The AOA noises are independent of the sensor position noise. The AOA noise power $\sigma_n^2 = (0.1^0)^2$ and the sensor positions is $\sigma_m^2 \times R$, where σ_m^2 varies from 10^{-4} to $10^2 \mathrm{m}^2$, $\boldsymbol{R} = \mathrm{diag}[1,1,2,2,10,10,40,40,20,20,100,100]$.

Fig. 3 (a) and (b) show the convergence performance of the logarithms over a range of initial guesses when the source location is in the near-field [650, 600]. Black areas in Fig. 3 represent initial guesses that do not converge to the global minimum, while the light colored regions show areas that converged. It could be concluded that Gauss-Newton is not a reliable method for AOAs localization. There are about 58.5% of initial guesses converged from Guass-Newton method, while the proposed algorithm is 99.8%. Fig. 3(c) and (d) are source location in the far-field [1500; 2500]. It is easy to find that the convergence region of Gauss-Newton is 55.7%, while the proposed method is 98.02%.

Fig. 4 shows a contour plot of cost function $\boldsymbol{J}_{\mathrm{NLS}}$ evaluated in a large region surrounding the true emitter position, which is located at [650; 600]. According to Fig. 4(a), the $\boldsymbol{J}_{\mathrm{NLS}}$ is nearly flat in a large neighborhood of the true source location, which presents convergence challenges for iterative methods. In order to prevent converging to the local minimum, the stopping criterion must be very strict. From Fig. 4(b), it can be seen that $\log 10(\boldsymbol{J}_{\mathrm{NLS}})$ is sufficiently small when reaching the global minimum [650; 600]. This property of AOAs localization could be used to indicate whether the iteration steps reach the global or local minimum.

Fig. 5 illustrates the test cases when the initial estimates are good or poor. In the good initial situation (Figs(a) and (c)), all algorithms converge to the source solution within a few steps when choosing the same stopping criterion. The convergence process of the standard LM method is slower than other algorithms, and the accelerated LM method needs the same steps as the Gauss-Newton method when the radius $r = 100$. In poor initial case (Figs(b) and (d)), the conventional Gauss-Newton algorithm oscillates at the beginning and divergences at the end, while the accelerated LM still performs well when $r = 100$. It is easy to get that the new algorithm has an increased convergence range and can cope with stronger non-linearity but requires less iteration steps than the standard LM and the Guass-Newton method.

Fig. 6 compares the MSE of the source position of the standard LM method, the proposed method and the CRLB. The CRLB is shown by the solid line, the square symbol is the MSE of the source location for standard LM method, and the cross symbol is for the accelerated LM method. It can be seen from the figure that both of the standard LM and the accelerated LM

method reach the CRLB for the source position. Actually, these iterative methods such as Gauss-Newton method and LM method have the same location accuracy when converging to the global minima.

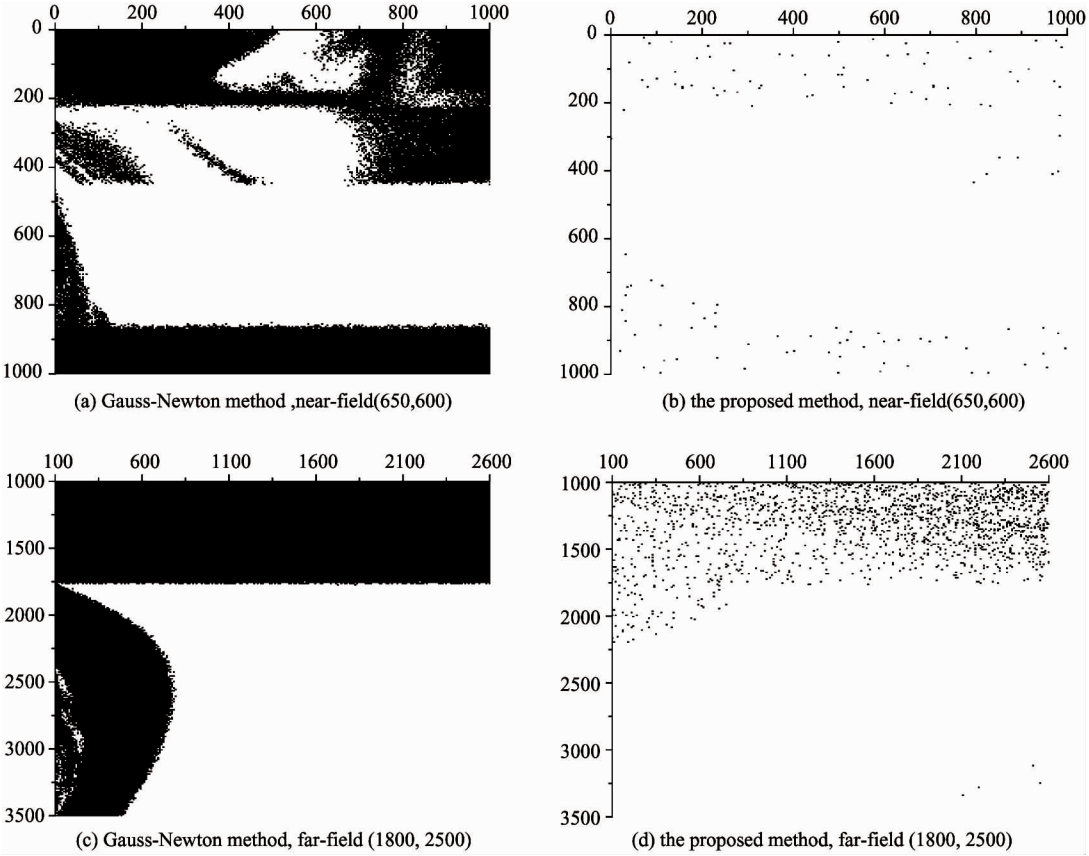

Fig. 3 Plot of initial guesses which converge for Gauss-Newton and the proposed method

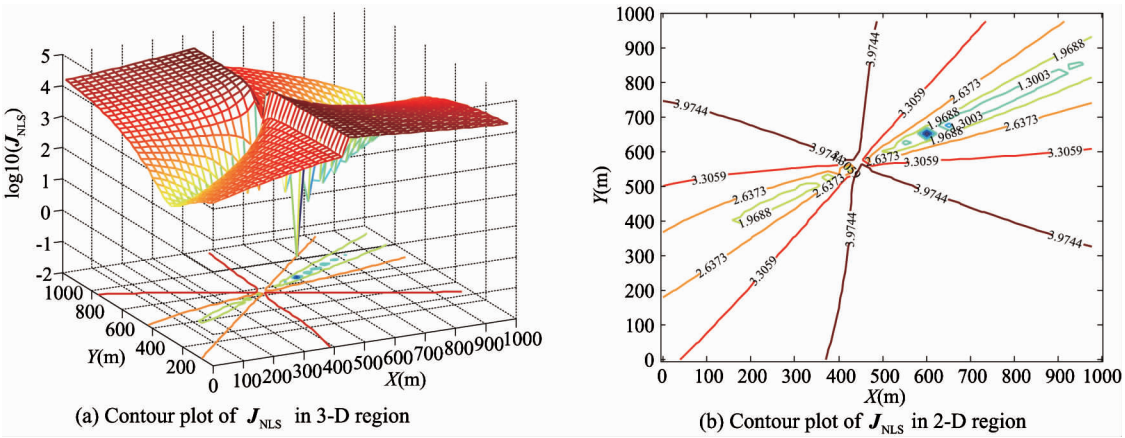

Fig. 4 Contour plot of J_{NLS} in Eq. (7) in large region surrounding true emitter position [650, 600]

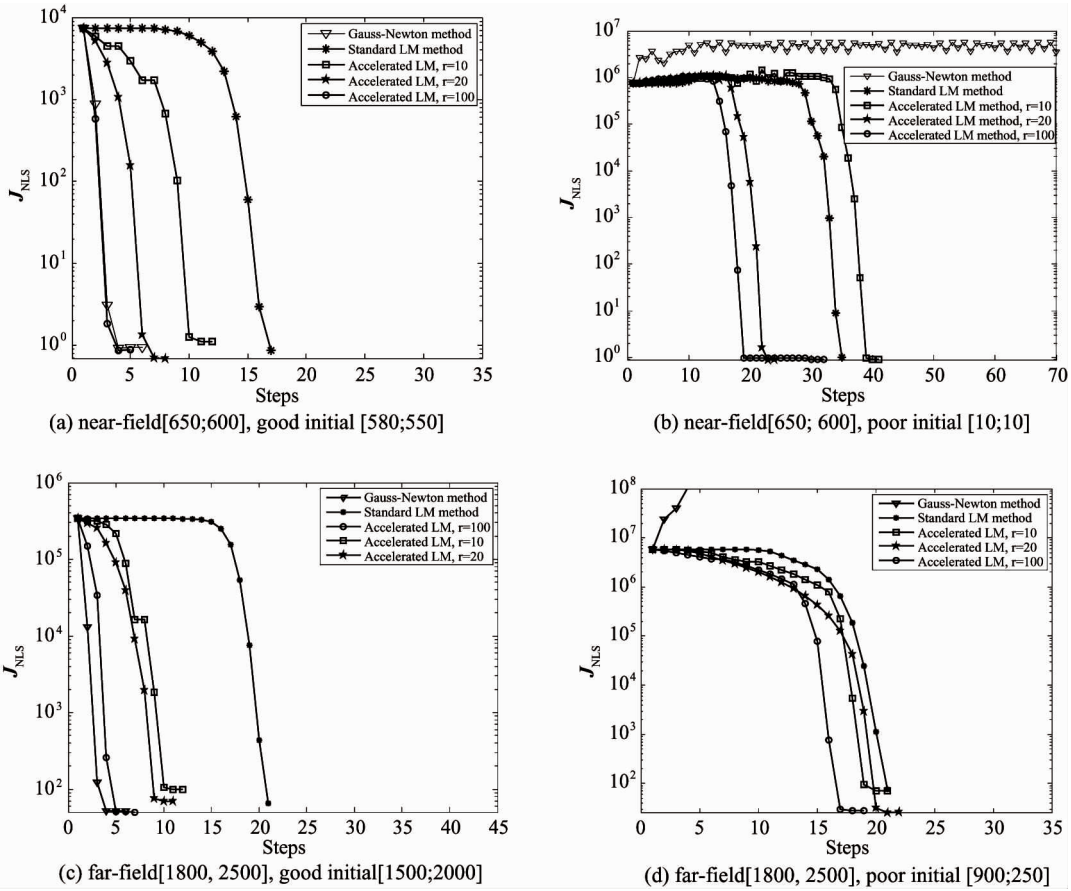

Fig. 5 Comparison of iteration convergence for AoAs localization

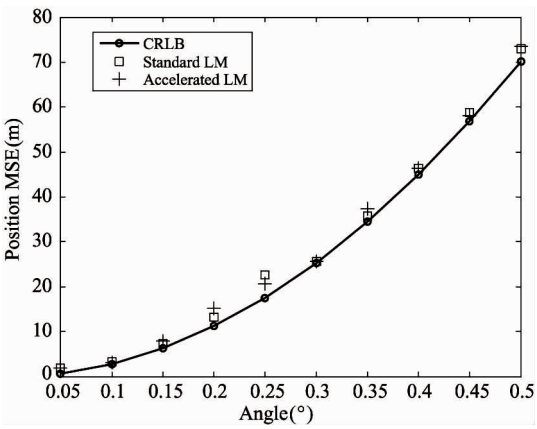

Fig. 6 Accuracy of source location estimates with σ_n varying from 0.05° to 0.5°

4 Conclusion

This paper presents a modified LM algorithm for the source localization. The method is suited for passive geolocation with AOAs measurement errors and sensor position errors, but not requires good prior knowledge of the emitter location. In order to increase the convergence rate, the trust region method is proposed to optimize the iteration procedure. Simulation results show that the algorithm converges to the global minimum from a wider initial guess, and the number of iteration steps is small.

References

[1] Hatem H, Kutluyil D. Passive localization of scanning emitters. *IEEE Transactions on Aerospace and electronic systems*, 2010, 46(2): 944-951

[2] Ho K C, Yang L. On the use of a calibration emitter for source localization in the presence of sensor position uncertainty. *IEEE transactions on signal procession*, 2008, 56(12): 5758-5772

[3] Lu X N, Ho K C. Taylor-series technique for source localization using AOAs in the presence of sensor location errors. In: Proceedings of the 4th IEEE Workshop on Sensor Array and Multichannel Processing, Waltham, USA, 2006. 190-194

[4] David F, Willett P, Bar-Shalom Y. CS versus MAP and MMOSPA for multi-target radar AOAs Svensson. In: Proceedings of the Conference Record of the 45th Asilomar Conference on Signals, Systems and Computers Lennart, Pacific Grove, USA, 2011. 1484-1490

[5] Lu L, Wu H C, Yan K. Robust expectation-maximization algorithm for multiple wideband acoustic source localization in the presence of nonuniform noise variances. *IEEE Sensors Journal*, 2011, 11(3):536-544

[6] Ho K C, Sun Ming. Passive source localization using time differences of arrival and gain ratios of arrival. *IEEE Transactions on Signal processing*, 2008, 56(2): 464-477

[7] Xie Y Q, Zhu P C. Grid-search-based hybrid TOA/AOA location techniques for NLOS environments. *IEEE communications letters*, 2009, 13(4): 254 -256

[8] Kwak Y, Hwang J. A new damping strategy of Levenberg-Marquardt algorithm for multilayer perceptrons. *Neural Network World*, 2011, 21(4): 327-340

[9] Behling R, Fischer A. A unified local convergence analysis of inexact constrained Levenberg-Marquardt methods. *Optimization Letters*, 2012, 6(5): 927-940

[10] Julio C, Sergiyenko O. Surface recognition improvement in 3D medical laser scanner using Levenberg-Marquardt method. *Signal Processing*, 2013, 93(2): 378-386

[11] Madsen K, Nielsen H B, Tingleff O. Non-linear least squares problems. Lyngby: Technical University of Denmark, 2004 : 33-54

Wu Xinhui, born in 1986. He received the B. S. degree in communication engineering from Huazhong University of Science and Technology in 2007, and the M. E. degree from Naval University of Engineering, Wuhan, China, in 2009. He is currently working towards the Ph. D. degree in communication and information engineering at Naval University of Engineering. His current research interests in the area of information fusion, passive target tracking and wireless communications.